

Docket No.: S&ZFH020501

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date indicated below.

By: Markus Nolff Date: January 16, 2004

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applic. No. : 10/713,691  
Applicant : Jürgen Herre, et al.  
Filed : November 14, 2003

Docket No. : S&ZFH020501  
Customer No. : 24131

CLAIM FOR PRIORITY

Commissioner for Patents,  
P.O. Box 1450, Alexandria, VA 22313-1450

Sir:

Claim is hereby made for a right of priority under Title 35, U.S. Code, Section 119, based upon the German Patent Application 101 23 281.0, filed May 14, 2001.

A certified copy of the above-mentioned foreign patent application is being submitted herewith.

Respectfully submitted,

Markus Nolff  
For Applicant

MARKUS NOLFF  
REG. NO. 37,006

Date: January 16, 2004

Lerner and Greenberg, P.A.  
Post Office Box 2480  
Hollywood, FL 33022-2480  
Tel: (954) 925-1100  
Fax: (954) 925-1101

/av

# BUNDESREPUBLIK DEUTSCHLAND



## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 101 23 281.0  
**Anmeldetag:** 14. Mai 2001  
**Anmelder/Inhaber:** Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung eV, München/DE  
**Bezeichnung:** Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion  
**IPC:** G 10 K 15/00

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.**

München, den 17. Dezember 2003  
**Deutsches Patent- und Markenamt**  
**Der Präsident**  
Im Auftrag

A handwritten signature in black ink, appearing to read "Agurke".

Agurke

PATENTANWÄLTE

European Patent Attorneys  
European Trademark Attorneys

Patentanwälte · Postfach 710867 · 81458 München

**Fraunhofer-Gesellschaft  
zur Förderung der  
angewandten Forschung e. v.  
Leonrodstraße 54  
80636 München**

Fritz Schoppe, Dipl.-Ing.  
Tankred Zimmermann, Dipl.-Ing.  
Ferdinand Stöckeler, Dipl.-Ing.  
Franz Zinkler, Dipl.-Ing.

Telefon/Telephone 089/790445-0  
Telefax/Facsimile 089/790 22 15  
Telefax/Facsimile 089/74996977  
e-mail: szsz\_iplaw@t-online.de

---

**Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von  
Rhythmusinformationen des Audiosignals unter Verwendung einer  
Autokorrelationsfunktion**

---

**Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion**

Beschreibung

Die vorliegende Erfindung bezieht sich auf Signalverarbeitungskonzepte und insbesondere auf die Analyse von Audiosignalen hinsichtlich von Rhythmusinformationen.

In den letzten Jahren ist die Verfügbarkeit multimedialen Datenmaterials, wie z. B. Audio- oder Video-Daten, stark gestiegen. Dies ist auf eine Reihe von technischen Faktoren zurückzuführen, welche sich insbesondere auf die breite Verfügbarkeit des Internets, von leistungsfähiger Rechnerhard- und Software sowie von leistungsfähigen Verfahren zur Datenkompression, d. h. Quellcodierung, von Audio- und Videoverfahren stützen.

Die riesigen Mengen audiovisueller Daten, die beispielsweise auf dem Internet weltweit verfügbar sind, verlangen nach Konzepten, die es ermöglichen, diese Daten nach inhaltlichen Kriterien beurteilen, katalogisieren usw. zu können. Es besteht der Wunsch, in der Lage zu sein, multimediale Daten gezielt durch Angabe sinnvoller Kriterien zu suchen und aufzufinden.

Dies erfordert sogenannte „inhaltsbasierte“ Techniken, die aus den audiovisuellen Daten sogenannte Merkmale, die in Fachkreisen auch als „Features“ bezeichnet werden, extrahieren, welche wichtige charakteristische Eigenschaften des Signals repräsentieren. Basierend auf solchen Merkmalen bzw. Kombinationen dieser Merkmale können Ähnlichkeitsbeziehungen bzw. Gemeinsamkeiten zwischen Audio- oder Videosignalen hergeleitet werden. Dieser Vorgang erfolgt durch Vergleich bzw. In-Beziehung-Setzen der extrahierten Merkmalswerte aus den verschiedenen Signalen, welche auch einfach als „Stücke“ bezeichnet werden.

Von besonderem Interesse ist die Bestimmung bzw. Extraktion von Merkmalen, die nicht nur signaltheoretische, sondern möglichst unmittelbare semantische Bedeutung haben, d. h. vom Hörer unmittelbar empfundene Eigenschaften repräsentieren.

Dies ermöglicht es dem Benutzer, auf einfache und intuitive Weise Suchanfragen zu formulieren, um Stücke aus dem gesamten vorhandenen Datenbestand einer Audiosignal-Datenbank zu finden. Ebenso erlauben es semantisch relevante Merkmale, Ähnlichkeitsbeziehungen zwischen Stücken zu modellieren, die der menschlichen Empfindung nahe kommen. Die Verwendung von Merkmalen, die semantische Bedeutung haben, ermöglicht auch beispielsweise ein automatisches Vorschlagen von für einen bestimmten Benutzer interessanten Stücken, wenn seine Vorlieben bekannt sind.

Im Bereich der Musikanalyse ist das Tempo ein wichtiger musikalischer Parameter, der semantische Bedeutung hat. Das Tempo wird üblicherweise in „Beats per Minute“ (BPM) gemessen. Die automatische Extraktion des Tempos sowie der Taktschwerpunkte des „Beats“ bzw. allgemein gesagt die automatische Extraktion von Rhythmusinformationen, ist ein Beispiel für die Gewinnung eines semantisch wichtigen Merkmals eines Musikstücks.

Weiterhin besteht ein Wunsch dahingehend, dass die Merkmalsextraktion, d. h. das Extrahieren von Rhythmusinformationen aus einem Audiosignal, robust und recheneffizient stattfinden kann. Robustheit bedeutet, dass es keine Rolle spielen darf, ob das Stück quellcodiert und wieder decodiert worden ist, ob das Stück über einen Lautsprecher abgespielt und von einem Mikrophon empfangen worden ist, oder ob es von einem Instrument oder einer Mehrzahl von Instrumenten gespielt wird.

Für die Bestimmung der Taktschwerpunkte und damit auch des Tempos, d. h. für die Bestimmung von Rhythmusinformationen, hat

sich in den Fachkreisen auch der Begriff „Beat Tracking“ etabliert. Aus dem Stand der Technik ist es bereits bekannt, ein Beat Tracking auf der Basis einer notenartigen bzw. transkribierten Signaldarstellung, z. B. im Midi-Format, durchzuführen. Ziel ist es jedoch, keine solchen Metadarstellungen zu benötigen, sondern eine Analyse direkt mit einem z. B. PCM-codierten oder allgemein gesagt digital vorliegenden Audiosignal vorzunehmen.

Die Fachveröffentlichung „Tempo and Beat Analysis of Acoustic Musical Signals“ von Eric D. Scheirer, J. Acoust. Soc. Am. 103:1, (Jan 1998), Seiten 588 - 601, offenbart ein Verfahren zur automatischen Extraktion eines rhythmischen Pulses aus musikalischen Exzerpten. Das Eingangssignal wird mittels einer Filterbank in eine Reihe von Teilbändern aufgespalten, beispielsweise in 6 Teilbänder mit Übergangsfrequenzen von 200 Hz, 400 Hz, 800 Hz, 1600 Hz und 3200 Hz. Für das erste Teilband wird eine Tiefpassfilterung durchgeführt. Für das letzte Teilband wird eine Hochpassfilterung durchgeführt, für die restlichen, dazwischen liegenden Teilbände wird eine Bandpassfilterung beschrieben. Jedes Teilband wird folgendermaßen verarbeitet. Das Teilbandsignal wird zunächst gleichgerichtet. In anderen Worten ausgedrückt wird der Absolutbetrag der Abtastwerte bestimmt. Die resultierenden n Werte werden dann geglättet, beispielsweise mit einer Mittelung über ein geeignetes Fenster, um ein Hüllkurvensignal zu erhalten. Zur Senkung der Rechenkomplexität kann das Hüllkurvensignal unterabgetastet werden. Die Hüllkurvensignale werden differenziert, d. h. plötzliche Änderungen der Signalamplitude werden durch das Differenzierungsfilter bevorzugt weitergeleitet. Das Resultat wird dann auf nicht-negative Werte begrenzt. Jedes Hüllkurvensignal wird dann in eine Bank resonanter Filter, d. h. Oszillatoren, gegeben, die jeweils ein Filter für jeden Tempobereich enthalten, so dass das zum musikalischen Tempo passende Filter am stärksten angeregt wird. Für jedes Filter wird die Energie des Ausgangssignals als Maß für die Übereinstimmung

des Tempos des Eingangssignals mit dem zum Filter gehörigen Tempo berechnet. Die Energien für jedes Tempo werden schließlich über alle Teilbänder aufsummiert, wobei die größte Energiesumme das als Resultat gelieferte Tempo, d. h. die Rhythmusinformationen, kennzeichnet. Im Gegensatz zu Autokorrelationsverfahren wird als vorteilhaft herausgestellt, daß die Oszillatorbank auch mit Ausgangssignalen bei dem doppelten, dreifachen, etc. des Tempos oder auch bei rationalen Vielfachen (z. B.  $2/3$ ,  $4/3$ ) des Tempos auf einen Stimulus reagiert. Eine Autokorrelationsfunktion hat diese Eigenschaft nicht, sie liefert nur Ausgangssignale bei dem halbierten, gedrittelten, etc. Tempo.

Ein wesentlicher Nachteil dieses Verfahrens besteht in der großen Rechen- und Speicherkomplexität insbesondere zur Realisierung der großen Zahl von parallelschwingenden „Oszillatoren“, von denen letztlich nur ein einziger ausgewählt wird. Dies macht eine effiziente Implementierung beispielsweise für Echtzeitanwendungen nahezu unmöglich.

Die Fachveröffentlichung „Pulse Tracking with a Pitch Tracker“ von Eric D. Scheirer, Proc. 1997 Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk, NY, Oct 1997, beschreibt einen Vergleich des oben beschriebenen „Oszillatorkonzeptes“ mit einem Alternativkonzept, das auf der Verwendung von Autokorrelationsfunktionen zur Extraktion der Periodizität aus einem Audiosignal, d. h. der Rhythmusinformationen eines Signals, aufbaut. Ein Algorithmus zur Modellierung der menschlichen Tonhöhenwahrnehmung, d. h. des Pitch, wird für das „Beat Tracking“ eingesetzt.

Der bekannte Algorithmus ist in Fig. 3 als Blockdiagramm dargestellt. Das Audiosignal wird über einen Audioeingang 300 einer Analyse-Filterbank 302 zugeführt. Die Analyse-Filterbank erzeugt aus dem Audioeingang eine Anzahl  $n$  von Kanälen, d. h. von einzelnen Teilbandsignalen. Jedes Teilbandsignal enthält

einen bestimmten Bereich an Frequenzen des Audiosignals. Die Filter der Analyse-Filterbank sind so ausgewählt, dass sie die Selektionscharakteristik des menschlichen Innenohrs annähern. Eine solche Analyse-Filterbank wird auch als Gamma-Ton-Filterbank bezeichnet.

In den Einrichtungen 304a bis 304c werden die Rhythmus-Informationen jedes Teilbandsignals ausgewertet. Für jedes Eingangssignal wird zunächst ein hüllkurvenartiges Ausgangssignal berechnet (entsprechend einer sogenannten „Inner Hair Cell“-Verarbeitung im Ohr) und unterabgetastet. Aus diesem Resultat wird eine Autokorrelationsfunktion (AKF) berechnet, um die Periodizität des Signals als Funktion der Verzögerung, d. h. des „Lag“ zu erhalten.

Am Ausgang der Einrichtungen 304a bis 304c liegt dann für jedes Teilbandsignal eine Autokorrelationsfunktion vor, welche die Rhythmus-Informationen jedes Teilbandsignals darstellt.

Die einzelnen Autokorrelationsfunktionen der Teilbandsignale werden dann in einer Einrichtung 306 durch Summation kombiniert, um eine Summenautokorrelationsfunktion (SAKF) zu erhalten, welche Aspekte der Rhythmus-Informationen des Signals am Audio-Eingang 300 wiedergibt. Diese Informationen können an einem Tempo-Ausgang 308 ausgegeben werden. Große Werte in der Summenautokorrelation zeigen an, dass für eine einer Spitze der SAKF zugeordneten Verzögerung (Lag) eine hohe Periodizität der Notenanfänge vorliegt. Daher wird beispielsweise der größte Wert der Summenautokorrelationsfunktion innerhalb der musikalisch sinnvollen Verzögerungen gesucht.

Musikalisch sinnvolle Verzögerungen sind beispielsweise der Tempobereich zwischen 60 bpm und 200 bpm. Die Einrichtung 306 kann ferner angeordnet sein, um eine Verzögerungszeit in Tempoinformationen umzusetzen. So entspricht beispielsweise eine Spalte

einer Verzögerung von einer Sekunde einem Tempo von 60 Schlägen pro Minute. Kleinere Verzögerungen deuten auf höhere Tempos hin, während größere Verzögerungen auf kleinere Tempos als 60 bpm hinweisen.

Dieses Verfahren hat gegenüber dem zuerst genannten Verfahren einen Vorteil dahingehend, dass keine Oszillatoren mit großem Rechen- und Speicheraufwand implementiert werden müssen. Andererseits ist das Konzept dahingehend nachteilhaft, dass die Qualität der Ergebnisse sehr stark von der Art des Audiosignals abhängt. Ist aus einem Audiosignal beispielsweise ein dominantes Rhythmusinstrument herauszuhören, so wird das in Fig. 3 beschriebene Konzept gut funktionieren. Ist dagegen die Stimme dominant, welche keine besonders eindeutigen Rhythmusinformationen liefern wird, so wird die Rhythmusbestimmung mehrdeutig sein. In dem Audiosignal könnte durchaus auch ein Band vorkommen, das lediglich Rhythmusinformationen enthält, wie z. B. ein höheres Frequenzband, in dem beispielsweise ein Hihat eines Schlagzeugs positioniert ist, oder ein niedriges Frequenzband, in dem die große Trommel eines Schlagzeugs auf der Frequenzskala positioniert ist. Aufgrund der Kombination der Einzelinformationen werden jedoch die einigermaßen eindeutigen Informationen dieser speziellen Teilbänder von den mehrdeutigen Informationen der anderen Teilbänder überlagert bzw. „verwässert“.

Ein weiteres Problem bei der Verwendung von Autokorrelationsfunktionen zum Extrahieren der Periodizität eines Teilbandsignals besteht darin, dass die Summen-Autokorrelationsfunktion, die durch die Einrichtung 306 gewonnen wird, mehrdeutig ist. Die Summen-Autokorrelationsfunktion am Ausgang 306 ist dahingehend mehrdeutig, dass auch beim Vielfachen einer Verzögerung eine Autokorrelationsfunktions-Spitze erzeugt wird. Dies ist daraus verständlich, dass eine Sinuskomponente mit einer Periode von  $t_0$ , wenn sie einer Autokorrelationsfunktionverarbeitung unterzo-

gen wird, neben dem gewünschten Maximum bei  $t_0$  auch Maxima bei Vielfachen der Verzögerungen, d. h. bei  $2t_0$ ,  $3t_0$ , usw. erzeugt.

Die Fachveröffentlichung „A Computationally Efficient Multipitch Analysis Model“, von Tolonen und Karjalainen, IEEE Transactions on Speech and Audio Processing, Band 8, Nr. 6, Nov. 2000, offenbart ein rechenzeiteffizientes Modell für eine Periodizitätsanalyse von komplexen Audiosignalen. Das Rechenmodell teilt das Signal in zwei Kanäle auf, und zwar in einen Kanal unter 1000 Hz und einen Kanal über 1000 Hz. Hieraus wird eine Autokorrelation des unteren Kanals und eine Autokorrelation der Hüllkurve des oberen Kanals berechnet. Schließlich werden die beiden Autokorrelationsfunktionen summiert. Um die Mehrdeutigkeiten der Summenautokorrelationsfunktion zu eliminieren, wird die Summenautokorrelationsfunktion weiterverarbeitet, um eine sogenannte Enhanced Summary Autocorrelation Function (ESACF) (Weiterentwickelte Summenautokorrelationsfunktion) zu erhalten. Diese Nachbearbeitung der Summenautokorrelationsfunktion beinhaltet ein wiederholtes Abziehen von mit ganzzahligen Faktoren gespreizten Versionen der Autokorrelationsfunktion von der Summenautokorrelationsfunktion mit nachfolgender Begrenzung auf nicht-negative Werte.

Nachteilig an diesem Konzept ist die Tatsache, dass die durch die Autokorrelationsfunktionen in den Teilbändern pro Teilband erhaltenen Mehrdeutigkeiten erst in der Summenautokorrelationsfunktion eliminiert werden, jedoch nicht unmittelbar dort, wo sie auftreten, nämlich in den einzelnen Teilbändern.

Ein weiterer Nachteil dieses Konzepts ist die Tatsache, daß die Autokorrelationsfunktion an sich keinen Hinweis auf das doppelte, dreifache, ... des Tempos liefert, dem eine Autokorrelationsspitze zugeordnet ist.

Die Aufgabe der vorliegenden Erfindung besteht darin, eine Vorrichtung und ein Verfahren zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen unter Verwendung einer Autokorrelationsfunktion zu schaffen, das robust und rechenzeiteffizient ist.

Diese Aufgabe wird durch eine Vorrichtung zum Analysieren eines Audiosignals nach Patentanspruch 1, 8 oder 9, oder durch ein Verfahren zum Analysieren eines Audiosignals nach Patentanspruch 7, 10, 11 gelöst.

Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, daß eine Nachbearbeitung einer Autokorrelationsfunktion teilbandweise durchgeführt werden kann, um die Mehrdeutigkeiten der Autokorrelationsfunktion für periodische Signale zu eliminieren bzw. Tempoinformationen, die eine Autokorrelationsverarbeitung nicht liefert, den durch eine Autokorrelationsfunktion erhaltenen Informationen hinzugefügt werden. Gemäß einem Aspekt der vorliegenden Erfindung wird eine Autokorrelationsfunktion-Nachbearbeitung der Teilbandsignale verwendet, um die Mehrdeutigkeiten bereits „an der Wurzel“ zu beseitigen, bzw. „fehlende“ Rhythmusinformationen hinzuzufügen.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird eine Nachbearbeitung der Summenautokorrelationsfunktion durchgeführt, um nachbearbeitete Rhythmus-Rohinformationen für das Audiosignal zu erhalten, so daß in den nachbearbeiteten Rhythmus-Rohinformationen ein Signalanteil bei einem ganzzahligen Bruchteil einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, hinzugefügt ist. Hierdurch ist es möglich, die durch eine Autokorrelationsfunktion nicht erhaltenen Rhythmusinformationen bei doppelten, dreifachen etc. Tempi bzw. bei rationalen Vielfachen durch Berechnen von um einen ganzzahligen Faktor oder um einen rationalen Faktor gestauchten Versionen der Autokorrelationsfunktion und durch Addieren dieser Versionen zur

ursprünglichen Autokorrelationsfunktion zu erzeugen. Im Gegensatz zum Stand der Technik, bei dem hierzu eine aufwendige Oszillatorenbank benötigt wird, geschieht dies erfahrungsgemäß mit einfach zu implementierenden Gewichtungs- und Additionsroutinen.

Gemäß einem weiteren Aspekt der vorliegenden Erfundung wird die Summenautokorrelationsfunktion ferner nachbearbeitet, indem eine mit einem Faktor, der größer als Null und kleiner als Eins ist, gewichtete, um einen ganzzahligen Faktor größer als Eins gespreizte Version der Rhythmus-Rohinformationen zu der Autokorrelationsfunktion subtrahiert wird. Dies hat den Vorteil einer Elimination der AKF-Mehrdeutigkeiten bei den ganzzahligen Vielfachen der Verzögerung, der eine Autokorrelationsspitze zugeordnet ist. Während im Stand der Technik keine Gewichtung der gespreizten Versionen der Autokorrelationsfunktion vor der Subtraktion durchgeführt wird, und eine Elimination der Mehrdeutigkeiten somit nur im theoretisch optimalen Fall erreicht wird, bei dem der Rhythmus sich ideal zyklisch wiederholt, liefert die gewichtete Subtraktion die Möglichkeit, durch geeignete Wahl der Gewichtungsfaktoren, die beispielsweise empirisch erfolgen kann, Rhythmusinformationen, die sich nicht ideal zyklisch wiederholen, zu berücksichtigen.

Gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfundung wird eine Autokorrelationsfunktion-Nachverarbeitung durchgeführt, indem die mittels einer Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen mit gestauchten und/oder gespreizten Versionen derselben kombiniert werden. Im Falle der Verwendung von gespreizten Versionen der Rhythmus-Rohinformationen werden die gespreizten Versionen von den Rhythmus-Rohinformationen subtrahiert, während im Falle von um ganzzählige Faktoren gestauchten Versionen der Autokorrelationsfunktion diese gestauchten Versionen zu den Rhythmus-Rohinformationen hinzugefügt werden.

Bei einem bevorzugten Ausführungsbeispiel der Erfindung wird die gestauchte/gespreizte Version vor dem Addieren bzw. Subtrahieren mit einem Faktor zwischen Null und Eins gewichtet.

Gemäß einem weiteren bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird eine Qualitätsbeurteilung der Rhythmus-Rohinformationen, um ein Signifikanzmaß zu erhalten, auf der Basis der nachverarbeiteten Rhythmus-Rohinformationen durchgeführt, derart, daß die Qualitätsbeurteilung nicht mehr durch Autokorrelationsfunktions-Artefakte beeinflußt wird. Damit wird eine sichere Qualitätsbeurteilung möglich, wodurch die Robustheit des Bestimmens von Rhythmusinformationen des Audiosignals weiter gesteigert werden kann.

Alternativ kann die Qualitätsbeurteilung bereits vor der AKF-Nachverarbeitung stattfinden. Dies hat den Vorteil, daß, wenn ein flacher Verlauf der Rhythmus-Rohinformationen festgestellt wird, d.h. keine ausgeprägten Rhythmusinformationen, auf die AKF-Nachverarbeitung für dieses Teilbandsignal verzichtet werden kann, da dieses Teilband aufgrund seiner wenig aussagekräftigen Rhythmus-Informationen beim Ermitteln der Rhythmusinformationen des Audiosignals ohnehin keine Rolle spielen wird. Auf diese Art und Weise kann der Rechen- und Speicheraufwand weiter reduziert werden.

In den einzelnen Frequenzbändern, d. h. den Teilbändern, liegen oftmals unterschiedlich günstige Bedingungen für das Auffinden von rhythmischen Periodizitäten vor. Während beispielsweise bei Popmusik oftmals im Bereich der Mitte, beispielsweise um 1 kHz, das Signal von nicht mit dem Beat korrespondierenden Gesang dominiert wird, sind in den höheren Frequenzbereichen oftmals vor allem Perkussionsklänge präsent, wie z. B. das Hihat des Schlagzeugs, welche eine sehr gute Extraktion rhythmischer Regelmäßigkeiten erlauben. Anders ausgedrückt beinhalten unterschiedliche Frequenzbänder je nach Audiosignal eine unterschiedliche Menge

an rhythmischer Information bzw. haben eine unterschiedliche Qualität oder Signifikanz für die Rhythmusinformationen des Audiosignals.

Das Audiosignal wird daher zunächst in Teilbandsignale zerlegt. Jedes Teilbandsignal wird hinsichtlich seiner Periodizität untersucht, um Rhythmus-Rohinformationen für jedes Teilbandsignal zu erhalten. Hierauf wird gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung eine Bewertung der Qualität der Periodizität jedes Teilbandsignals durchgeführt, um ein Signifikanzmaß für jedes Teilbandsignal zu erhalten. Ein hohes Signifikanzmaß deutet darauf hin, dass in diesem Teilbandsignal deutliche Rhythmusinformationen vorliegen, während ein niedriges Signifikanzmaß darauf hindeutet, dass in diesem Teilbandsignal weniger eindeutige Rhythmusinformationen vorliegen.

Gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird bei der Untersuchung eines Teilbandsignals hinsichtlich seiner Periodizitäten zunächst eine modifizierte Hüllkurve des Teilbandsignals berechnet und dann eine Autokorrelationsfunktion der Hüllkurve berechnet. Die Autokorrelationsfunktion der Hüllkurve stellt die Rhythmus-Rohinformationen dar. Eindeutige Rhythmusinformationen liegen vor, wenn die Autokorrelationsfunktion deutliche Maxima aufweist, während weniger eindeutige Rhythmusinformationen vorliegen, wenn die Autokorrelationsfunktion der Hüllkurve des Teilbandsignals weniger ausgeprägte Signalspitzen oder überhaupt keine Signalspitzen hat. Eine Autokorrelationsfunktion, die deutliche Signalspitzen hat, wird daher ein hohes Signifikanzmaß erhalten, während eine Autokorrelationsfunktion, die einen relativ flachen Verlauf hat, ein niedriges Signifikanzmaß erhalten wird. Die Artefakte der Autokorrelationsfunktionen werden, wie oben ausgeführt, erfindungsgemäß eliminiert.

da Teilbandsignale mit keinen eindeutigen bzw. sogar abweichen-  
den Rhythmusinformationen, d. h. wenn der Gesang einen anderen  
Rhythmus aufweist als der eigentliche Beat des Stücks, die  
Rhythmusinformationen des Audiosignals nicht „verwässern“ bzw.  
„verfälschen“. Darüber hinaus werden sehr rauschartige Teilband-  
signale, welche eine Systemautokorrelationsfunktion mit voll-  
ständig flachem Verlauf liefern, das Signal/Rausch-Verhältnis  
bei der Bestimmung der Rhythmusinformationen nicht verschlech-  
tern. Genau dies würde jedoch auftreten, wenn, wie im Stand der  
Technik, einfach alle Autokorrelationsfunktionen der Teilbandsi-  
gnale mit gleichem Gewicht aufsummiert werden.

Ein weiterer Vorteil des beschriebenen Verfahrens besteht darin,  
dass mit einem kleinen zusätzlichen Rechenaufwand ein Signifi-  
kanzmaß bestimmt werden kann, und dass die Bewertung der Rhyth-  
mus-Rohinformationen mit dem Signifikanzmaß und die anschließen-  
de Summierung ohne großen Speicher- und Recheneitaufwand effi-  
zient durchgeführt werden können, was das erfindungsgemäße Kon-  
zept insbesondere auch für Echtzeitanwendungen empfiehlt.

Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung wer-  
den nachfolgend bezugnehmend auf die beiliegenden Zeichnungen  
detailliert erläutert. Es zeigen:

Fig. 1 ein Blockschaltbild einer Vorrichtung zum Analysieren  
eines Audiosignals mit einer Qualitätsbewertung der  
Rhythmus-Rohinformationen;

Fig. 2 ein Blockschaltbild einer Vorrichtung zum Analysieren  
eines Audiosignals unter Verwendung von Gewichtungsfak-  
toren auf der Basis der Signifikanzmaße;

Fig. 3 ein Blockschaltbild einer bekannten Vorrichtung zum Ana-  
lysieren eines Audiosignals hinsichtlich von Rhythmusin-  
formationen;

Fig. 4 ein Blockschaltbild einer Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen unter Verwendung einer Autokorrelationsfunktion mit einer teilbandweisen Nachbearbeitung der Rhythmus-Rohinformationen; und

Fig. 5 ein detailliertes Blockschaltbild der Einrichtung zum Nachbearbeiten von Fig. 4.

Fig. 1 zeigt ein Blockschaltbild einer Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen. Das Audiosignal wird über einen Eingang 100 einer Einrichtung 102 zum Zerlegen des Audiosignals in zumindest zwei Teilbandsignale 104a und 104b zugeführt. Jedes Teilbandsignal 104a, 104b wird in eine Einrichtung 106a bzw. 106b zum Untersuchen desselben hinsichtlich von Periodizitäten in dem Teilbandsignal zugeführt, um Rhythmus-Rohinformationen 108a bzw. 108b für jedes Teilbandsignal zu erhalten. Die Rhythmus-Rohinformationen werden dann einer Einrichtung 110a bzw. 110b zum Bewerten einer Qualität der Periodizität jedes der zumindest zwei Teilbandsignale zugeführt, um ein Signifikanzmaß 112a, 112b für jedes der zumindest zwei Teilbandsignale zu erhalten. Sowohl die Rhythmus-Rohinformationen 108a, 108b als auch die Signifikanzmaße 112a, 112b werden einer Einrichtung 114 zum Ermitteln der Rhythmusinformationen des Audiosignals zugeführt. Die Einrichtung 114 berücksichtigt beim Ermitteln der Rhythmusinformationen des Audiosignals die Signifikanzmaße 112a, 112b für die Teilbandsignale sowie die Rhythmus-Rohinformationen 108a, 108b von zumindest einem Teilbandsignal.

Hat die Einrichtung 110a zur Qualitätsbewertung beispielsweise festgestellt, dass in dem Teilbandsignal 104a keine besondere Periodizität vorliegt, so wird das Signifikanzmaß 112a sehr klein, bzw. gleich 0 sein. In diesem Fall stellt die Einrichtung

114 zum Ermitteln der Rhythmusinformationen fest, dass das Signifikanzmaß 112a gleich Null ist, so dass die Rhythmus-Rohinformationen 108a des Teilbandsignals 104a bei der Ermittlung der Rhythmusinformationen des Audiosignals überhaupt nicht mehr berücksichtigt werden müssen. Die Rhythmusinformationen des Audiosignals werden dann allein und ausschließlich auf der Basis der Rhythmus-Rohinformationen 108b des Teilbandsignals 104b bestimmt.

Im nachfolgenden wird auf Fig. 2 hinsichtlich einer speziellen Ausführungsform der Vorrichtung von Fig. 1 eingegangen. Als Einrichtung 102 zum Zerlegen des Audiosignals kann eine übliche Analyse-Filterbank verwendet werden, die ausgangsseitig eine von einem Benutzer wählbare Anzahl von Teilbandsignalen liefert. Jedes Teilbandsignal wird dann der Verarbeitung der Einrichtungen 106a, 106b bzw. 106c unterzogen, woraufhin dann durch die Einrichtungen 110a bis 110c von jeden Rhythmus-Rohinformationen Signifikanzmaße ermittelt werden. Die Einrichtung 114 umfaßt bei der in Fig. 2 dargestellten bevorzugten Ausführungsform eine Einrichtung 114a zum Berechnen von Gewichtungsfaktoren für jedes Teilbandsignal auf der Basis des Signifikanzmaßes für dieses Teilbandsignal und optional auch der anderen Teilbandsignale. In der Einrichtung 114b findet dann eine Gewichtung der Rhythmus-Rohinformationen 108a bis 108c mit dem Gewichtungsfaktor für dieses Teilbandsignal statt, woraufhin dann, ebenfalls in der Einrichtung 114b, die gewichteten Rhythmus-Rohinformationen kombiniert, z. B. aufsummiert, werden, um an dem Tempo-Ausgang 116 die Rhythmusinformationen des Audiosignals zu erhalten.

Das erfindungsgemäße Konzept stellt sich somit folgendermaßen dar. Nach der Evaluierung der rhythmischen Informationen der Einzelbänder, welche beispielsweise durch Hüllkurvenbildung, Glättung, Differenzieren, Begrenzen auf positive Werte und Bilden der Autokorrelationsfunktion stattfinden kann (Einrichtungen 106a bis 106c), findet eine Bewertung der Wertigkeit bzw. der

Qualität dieser Zwischenergebnisse in den Einrichtungen 110a bis 110c statt. Dies wird mit Hilfe einer Bewertungsfunktion erreicht, welche die Verlässlichkeit der jeweiligen Einzelergebnisse mit einem Signifikanzmaß bewertet. Aus den Signifikanzmaßen aller Teilbandsignale wird ein Gewichtungsfaktor für jedes Band für die Extraktion der Rhythmus-Informationen hergeleitet. Das Gesamtergebnis der Rhythmusextraktion wird dann in der Einrichtung 114b durch Kombination der bandweisen Einzelergebnisse unter Berücksichtigung ihrer jeweiligen Gewichtungsfaktoren erreicht.

Als Resultat zeigt ein solchermaßen implementierter Algorithmus zur Rhythmusanalyse eine gute Fähigkeit, rhythmische Informationen in einem Signal auch unter ungünstigen Bedingungen zuverlässig zu finden. Das erfindungsgemäße Konzept zeichnet sich daher durch eine hohe Robustheit aus.

Bei einer bevorzugten Ausführungsform werden die Rhythmus-Rohinformationen 108a, 108b, 108c, die die Periodizität des jeweiligen Teilbandsignals darstellen, mittels einer Autokorrelationsfunktion bestimmt. In diesem Fall wird es bevorzugt, das Signifikanzmaß zu bestimmen, indem ein Maximum der Autokorrelationsfunktion durch einen Mittelwert der Autokorrelationsfunktion geteilt wird, und dann der Wert 1 subtrahiert wird. Es sei darauf hingewiesen, daß jede Autokorrelationsfunktion immer bei einer Verzögerung von 0 ein lokales Maximum, d. h. eine Spitze, liefert, die die Energie des Signals repräsentiert. Dieses lokale Maximum sollte unberücksichtigt bleiben, damit die Qualitätsbestimmung nicht verfälscht wird.

Ferner soll die Autokorrelationsfunktion lediglich in einem speziellen Tempobereich betrachtet werden, d. h. von einer maximalen Verzögerung, die dem kleinsten interessierenden Tempo entspricht, zu einer minimalen Verzögerung, die dem höchsten inter-

essierenden Tempo entspricht. Ein typischer Tempobereich liegt zwischen 60 bpm und 200 bpm.

Alternativ kann als Signifikanzmaß das Verhältnis zwischen dem arithmetischen Mittelwert der Autokorrelationsfunktion im interessierenden Tempobereich und dem geometrischen Mittelwert der Autokorrelationsfunktion im interessierenden Tempobereich bestimmt werden. Es ist bekannt, dass wenn alle Werte der Autokorrelationsfunktion gleich sind, d. h. wenn die Autokorrelationsfunktion einen flachen Verlauf hat, der geometrische Mittelwert der Autokorrelationsfunktion und der arithmetische Mittelwert der Autokorrelationsfunktion gleich sind. In diesem Fall hätte das Signifikanzmaß einen Wert gleich 1, was bedeutet, dass die Rhythmus-Rohinformationen nicht signifikant sind.

Im Falle einer Systemautokorrelationsfunktion mit starken Spitzen würde das Verhältnis aus arithmetischem Mittelwert zu geometrischem Mittelwert größer 1 sein, was bedeutet, dass die Autokorrelationsfunktion gute Rhythmusinformationen aufweist. Je kleiner jedoch das Verhältnis zwischen arithmetischem Mittelwert und geometrischem Mittelwert wird, desto flacher ist die Autokorrelationsfunktion und desto weniger Periodizitäten enthält sie, was wiederum bedeutet, dass die Rhythmusinformationen dieses Teilbandsignals weniger signifikant, d. h. eine geringere Qualität haben, was sich in einem geringen bzw. einem Gewichtungsfaktor von 0 äußern wird.

Bezüglich der Gewichtungsfaktoren existieren verschiedene Möglichkeiten. Bevorzugt wird eine relative Gewichtung, derart, dass sich alle Gewichtungsfaktoren sämtlicher Teilbandsignale zu 1 aufaddieren, d. h. dass der Gewichtungsfaktor eines Bandes bestimmt wird als der Signifikanzwert dieses Bandes geteilt durch die Summe aller Signifikanzwerte. In diesem Fall wird eine relative Gewichtung vor der Aufsummation der gewichteten Rhythmus-

Rohinformationen durchgeführt, um die Rhythmus-Informationen des Audiosignals zu erhalten.

Wie bereits ausgeführt worden ist, wird es bevorzugt, die Auswertung der Rhythmus-Informationen unter Verwendung einer Autokorrelationsfunktion durchzuführen. Dieser Fall ist in Fig. 4 dargestellt. Das Audiosignal wird über den Audiosignaleingang 100 in die Einrichtung 102 zum Zerlegen des Audiosignals in Teilbandsignale 104a und 104b eingespeist. Jedes Teilbandsignal wird dann in der Einrichtung 106a bzw. 106b, wie es ausgeführt worden ist, unter Verwendung einer Autokorrelationsfunktion untersucht, um die Periodizität des Teilbandsignals zu ermitteln.

Am Ausgang der Einrichtung 106a bzw. 106b liegen dann die Rhythmus-Rohinformationen 108a, 108b vor. Diese werden in eine Einrichtung 118a bzw. 118b eingespeist, um die mittels der Autokorrelationsfunktion von der Einrichtung 116a ausgegebenen Rhythmus-Rohinformationen nachzubearbeiten. Damit wird u. a. sichergestellt, daß die Mehrdeutigkeiten der Autokorrelationsfunktion, d. h. daß bei ganzzahligen Vielfachen der Verzögerungen ebenfalls Signalspitzen auftreten, teilbandweise eliminiert werden, um nachbearbeitete Rhythmus-Rohinformationen 120a bzw. 120b zu erhalten.

Dies hat den Vorteil, dass die Mehrdeutigkeiten der Autokorrelationsfunktionen, d. h. der Rhythmus-Rohinformationen 108a, 108b, bereits teilbandweise eliminiert werden, und nicht erst, wie im Stand der Technik, nach der Summation der einzelnen Autokorrelationsfunktionen. Darüber hinaus ermöglicht die einzelbandweise Eliminierung der Mehrdeutigkeiten in den Autokorrelationsfunktionen durch die Einrichtungen 118a, 118b, dass die Rhythmus-Rohinformationen der Teilbandsignale unabhängig voneinander gehandhabt werden können. Sie können beispielsweise einer Qualitätsbeurteilung mittels der Einrichtung 110a für die Rhythmus-Rohinformationen 108a oder mittels der Einrichtung 110b für die Rhythmus-Rohinformationen 108b unterzogen werden.

Wie es durch die gestrichelten Linien in Fig. 4 dargestellt ist, kann die Qualitätsbeurteilung jedoch auch anhand der nachbearbeiteten Rhythmus-Rohinformationen stattfinden, wobei diese letztere Möglichkeit bevorzugt wird, da die Qualitätsbeurteilung auf der Basis der nachbearbeiteten Rhythmus-Rohinformationen sicherstellt, dass die Qualität einer Information beurteilt wird, die nicht mehr vieldeutig ist.

Die Ermittlung der Rhythmus-Informationen durch die Einrichtung 114 findet dann auf der Basis von nachbearbeiteten Rhythmus-Informationen eines Kanals und vorzugsweise auch auf der Basis des Signifikanzmaßes für diesen Kanal statt.

Wenn eine Qualitätsbeurteilung auf der Basis der Rhythmus-Rohinformationen, also des Signals vor der Einrichtung 118a durchgeführt wird, ist dies dahingehend vorteilhaft, dass, wenn bestimmt wird, dass das Signifikanzmaß gleich 0 ist, d. h. dass die Autokorrelationsfunktion einen flachen Verlauf hat, auf die Nachbearbeitung mittels der Einrichtung 118a gänzlich verzichtet werden kann, um Rechenzeitressourcen einzusparen.

Im nachfolgenden wird auf Fig. 5 eingegangen, um einen detaillierteren Aufbau einer Einrichtung 118a oder 118b zur Nachbearbeitung der Rhythmus-Rohinformationen darzustellen. Zunächst wird das Teilband-Signal beispielsweise 104a, in die Einrichtung 106a zur Untersuchung der Periodizität des Teilbandsignals mittels einer Autokorrelationsfunktion eingespeist, um Rhythmus-Rohinformationen 108a zu erhalten. Um die Mehrdeutigkeiten teilbandweise zu eliminieren, kann genauso wie im Stand der Technik, eine gespreizte Autokorrelationsfunktion mittels einer Einrichtung 121 berechnet werden, wobei die Einrichtung 121 angeordnet ist, um die gespreizte Autokorrelationsfunktion so zu berechnen, dass sie um ein ganzzahliges Vielfaches gespreizt ist. Eine Einrichtung 122 ist in diesem Fall angeordnet, um die gespreizte

Autokorrelationsfunktion von der ursprünglichen Autokorrelationsfunktion, d. h. den Rhythmus-Rohinformationen 108a zu subtrahieren. Insbesondere wird es bevorzugt, zunächst eine auf das Doppelte gespreizte Autokorrelationsfunktion in der Einrichtung 121 zu berechnen und dann von den Rhythmus-Rohinformationen 108a zu subtrahieren. Dann, in dem nächsten Schritt, wird eine um den Faktor 3 gespreizte Autokorrelationsfunktion in der Einrichtung 121 berechnet und von dem Ergebnis der vorherigen Subtraktion wieder abgezogen, so dass nach und nach sämtliche Mehrdeutigkeiten aus den Rhythmus-Rohinformationen eliminiert werden.

Darüber hinaus können die gespreizten Versionen der Rhythmus-Rohinformationen 108a vor dem Subtrahieren gewichtet werden, um auch hier eine Flexibilität im Sinne einer hohen Robustheit zu erreichen.

Durch das Verfahren, die Periodizität eines Teilbandsignals auf der Basis einer Autokorrelationsfunktion zu untersuchen, kann also eine weitere Verbesserung erzielt werden, wenn die Eigenschaften der Autokorrelationsfunktion mit einbezogen werden und die Nachbearbeitung unter Verwendung der Einrichtung 118a oder 118b durchgeführt wird. So erzeugt eine periodische Folge von Notenanfängen mit einem Abstand  $t_0$  nicht nur eine AKF-Spitze bei einer Verzögerung  $t_0$  sondern auch bei  $2t_0$ ,  $3t_0$ , etc. Dies wird zu einer Vieldeutigkeit in der Tempodetektion, d. h. dem Suchen signifikanter Maxima in der Autokorrelationsfunktion, führen. Die Mehrdeutigkeiten können dadurch eliminiert werden, wenn um ganzzahlige Faktoren gespreizte Versionen der AKF vom Ausgangswert teilbandweise (gewichtet) abgezogen werden.

Darüber hinaus können die gestauchten Versionen der Rhythmus-Rohinformationen 108a vor dem Addieren mit einem Faktor ungleich Eins gewichtet werden, um auch hier eine Flexibilität im Sinne einer hohen Robustheit zu erreichen.

Ferner besteht bei der Autokorrelationsfunktion das Problem, daß sie keine Informationen bei  $t_0/2$ ,  $t_0/3$  ... usw., also beim Doppelten, Dreifachen, etc. des „Grundtempo“ liefert, was besonders dann zu falschen Ergebnissen führen kann, wenn zwei Instrumente, die in unterschiedlichen Teilbändern liegen, zusammen den Rhythmus des Signals definieren. Diese Sache wird dadurch berücksichtigt, dass um ganzzahlige Faktoren gestauchte Versionen der Autokorrelationsfunktion berechnet werden und diese dann zu den Rhythmus-Rohinformationen gewichtet oder ungewichtet hinzugefügt werden.

Die AKF-Nachverarbeitung findet somit teilbandweise statt, wobei für mindestens ein Teilbandsignal eine Autokorrelationsfunktion errechnet wird und diese mit gedehnten oder gespreizten Versionen dieser Funktion kombiniert wird.

Gemäß einem anderen Aspekt der vorliegenden Erfindung wird zunächst die Summenautokorrelationsfunktion der Teilbänder erzeugt, woraufhin um ganzzahligen Faktoren gestauchte Versionen der Summenautokorrelationsfunktion vorzugsweise gewichtet hinzugefügt werden, um die Unzulänglichkeiten der Autokorrelationsfunktion bei dem doppelten, dreifachen etc. Tempo auszuräumen.

Gemäß einem weiteren Aspekt wird die Nachbearbeitung der Summenautokorrelationsfunktion, um die Mehrdeutigkeiten bei dem halben, dem dritten Teil, dem vierten Teil etc. des Tempos zu eliminieren, durchgeführt, indem die um ganzzahlige Faktoren gespreizten Versionen der Summenautokorrelationsfunktion nicht einfach subtrahiert werden, sondern vor der Subtraktion mit einem Faktor ungleich Eins und vorzugsweise kleiner als Eins und größer als Null gewichtet werden und dann erst subtrahiert werden. Damit wird eine robustere Bestimmung der Rhythmusinformationen möglich, da das ungewichtete Subtrahieren lediglich für

ideale sinusförmige Signale eine vollständige Elimination der AKF-Mehrdeutigkeiten liefert.

Patentansprüche

1. Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Merkmalen:

einer Einrichtung (102) zum Zerlegen des Audiosignals in zumindest zwei Teilbandsignale (104a, 104b);

einer Einrichtung zum Untersuchen (106a, 106b) zumindest eines Teilbandsignals hinsichtlich einer Periodizität in dem zumindest einen Teilbandsignal mittels einer Autokorrelationsfunktion, um Rhythmus-Rohinformationen (108a) für das Teilbandsignal zu erhalten, wobei einer Spurze der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

einer Einrichtung (118a) zum Nachbearbeiten der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen (108a) für das Teilbandsignal (104a), um nachbearbeitete Rhythmus-Rohinformationen (120a) für das Teilbandsignal zu erhalten, so daß in den nachbearbeiteten Rhythmus-Rohinformationen eine Mehrdeutigkeit bei einem ganzzahligen Vielfachen einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, vermindert ist, oder ein Signalanteil bei einem ganzzahligen Bruchteil einer Verzögerung, der eine Autokorrelationsfunktionspitze zugeordnet ist, hinzugefügt ist; und

einer Einrichtung (114) zum Ermitteln der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen (120a) des Teilbandsignals und unter Verwendung eines weiteren Teilbandsignals der zumindest zwei Teilbandsignale.

2. Vorrichtung nach Anspruch 1,

bei der die Einrichtung zum Nachbearbeiten (118a, 118b) folgende Merkmale aufweist:

eine Einrichtung (121) zum Berechnen einer um einen ganzzahligen Faktor gespreizten Version der Rhythmus-Rohinformationen (108a) eines Teilbandsignals; und

einer Einrichtung (122) zum Subtrahieren der um einen ganzzahligen Faktor größer als Eins gespreizten Version der Rhythmus-Rohinformationen (108a) des Teilbandsignals oder eine von dieser Version abgeleitete Version von den Rhythmus-Rohinformationen (108a) des Teilbandsignals, um die nachbearbeiteten Rhythmus-Rohinformationen (120a) für das Teilbandsignal zu erhalten.

3. Vorrichtung nach Anspruch 2, bei der die Einrichtung (122) zum Subtrahieren angeordnet ist, um vor dem Subtrahieren eine Gewichtung der gespreizten Version mit einem Faktor zwischen Null und Eins durchzuführen, um die abgeleitete Version zu erzeugen.

4. Vorrichtung nach Anspruch 1, bei der die Einrichtung zum Nachbearbeiten (118a) folgende Merkmale aufweist:

eine Einrichtung (121) zum Berechnen einer um einen ganzzahligen Faktor größer als Eins gestauchten Version der Rhythmus-Rohinformationen (108a); und

eine Einrichtung (122) zum Addieren der gestauchten Version der Rhythmus-Rohinformationen des Teilbandsignals oder eine von dieser Version abgeleitete Version zu den Rhythmus-Rohinformationen (108a) des Teilbandsignals, um die nachbearbeiteten Rhythmus-Rohinformationen (120a) für das Teilbandsignal zu erhalten.

5. Vorrichtung nach Anspruch 4, bei der die Einrichtung (122) zum Addieren angeordnet ist, um vor der Addition eine Gewichtung

der gestauchten Version der Rhythmus-Rohinformationen mittels eines Faktors zwischen Null und Eins durchzuführen, derart, daß eine gewichtete gestauchte Version der Rhythmus-Rohinformationen zu den Rhythmus-Rohinformationen des Teilbandsignals hinzugefügt wird, um die abgeleitete Version zu erzeugen.

6. Vorrichtung nach einem der vorhergehenden Ansprüche, die ferner folgendes Merkmal aufweist:

eine Einrichtung (110a, 110b) zum Bewerten einer Qualität der Periodizität der nachverarbeiteten Rhythmus-Rohinformationen (120a), um ein Signifikanzmaß für das Teilbandsignal zu erhalten,

wobei die Einrichtung (114) zum Ermitteln ferner angeordnet ist, um die Rhythmusinformationen des Audiosignals unter Berücksichtigung des Signifikanzmaßes des Teilbandsignals zu ermitteln.

7. Verfahren zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Schritten:

Zerlegen (102) des Audiosignals in zumindest zwei Teilbandsignale (104a, 104b);

Untersuchen (106a, 106b) zumindest eines Teilbandsignals hinsichtlich einer Periodizität in dem zumindest einen Teilbandsignal mittels einer Autokorrelationsfunktion, um Rhythmus-Rohinformationen (108a) für das Teilbandsignal zu erhalten, wobei einer Spitze der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

Nachbearbeiten (118a) der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen (108a) für das Teilbandsignal (104a), um nachbearbeitete Rhythmus-Rohinformationen (120a)

für das Teilbandsignal zu erhalten, so daß in den nachbearbeiteten Rhythmus-Rohinformationen eine Mehrdeutigkeit bei einem ganzzahligen Vielfachen einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, vermindert ist, oder ein Signalanteil bei einem ganzzahligen Bruchteil einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, hinzugefügt ist; und

=

Ermitteln (114) der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen (120a) des Teilbandsignals und unter Verwendung eines weiteren Teilbandsignals der zumindest zwei Teilbandsignale.

8. Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Merkmalen:

einer Einrichtung zum Untersuchen des Audiosignals hinsichtlich einer Periodizität in dem Audiosignal, um Rhythmus-Rohinformationen für das Audiosignal zu erhalten, wobei einer Spitze der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

einer Einrichtung zum Nachbearbeiten der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen für das Audiosignal, um nachbearbeitete Rhythmus-Rohinformationen für das Audiosignal zu erhalten, so daß in den nachbearbeiteten Rhythmus-Rohinformationen ein Signalanteil bei einem ganzzahligen Bruchteil einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, hinzugefügt ist; und

einer Einrichtung zum Ermitteln der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen des Audiosignals.

9. Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Merkmalen:

einer Einrichtung zum Untersuchen des Audiosignals hinsichtlich einer Periodizität in dem Audiosignal, um Rhythmus-Rohinformationen für das Audiosignal zu erhalten, wobei einer Spalte der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

einer Einrichtung zum Nachbearbeiten der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen für das Audiosignal, um nachbearbeitete Rhythmus-Rohinformationen für das Audiosignal zu erhalten, durch Subtrahieren einer mit einem Faktor ungleich Eins gewichteten, um einen ganzzahligen Faktor größer als Eins gespreizten Version der Rhythmus-Rohinformationen; und

einer Einrichtung zum Ermitteln der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen des Audiosignals.

10. Verfahren zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Schritten:

Untersuchen des Audiosignals hinsichtlich einer Periodizität in dem Audiosignal, um Rhythmus-Rohinformationen für das Audiosignal zu erhalten, wobei einer Spalte der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

Nachbearbeiten der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen für das Audiosignal, um nachbearbeitete Rhythmus-Rohinformationen für das Audiosignal zu erhalten, so daß in den nachbearbeiteten Rhythmus-Rohinformationen

ein Signalanteil bei einem ganzzahligen Bruchteil einer Verzögerung, der eine Autokorrelationsfunktionsspitze zugeordnet ist, hinzugefügt ist; und

Ermitteln der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen des Audiosignals.

11. Verfahren zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion, mit folgenden Schritten:

- Untersuchen des Audiosignals hinsichtlich einer Periodizität in dem Audiosignal, um Rhythmus-Rohinformationen für das Audiosignal zu erhalten, wobei einer Spitze der Autokorrelationsfunktion eine Verzögerung zugeordnet ist;

Nachbearbeiten der mittels der Autokorrelationsfunktion bestimmten Rhythmus-Rohinformationen für das Audiosignal, um nachbearbeitete Rhythmus-Rohinformationen für das Audiosignal zu erhalten, durch Subtrahieren einer mit einem Faktor ungleich Eins gewichteten, um einen ganzzahligen Faktor größer als Eins gespreizten Version der Rhythmus-Rohinformationen; und

- Ermitteln der Rhythmusinformationen des Audiosignals unter Verwendung der nachbearbeiteten Rhythmus-Rohinformationen des Audiosignals.

**Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von  
Rhythmusinformationen des Audiosignals unter Verwendung einer  
Autokorrelationsfunktion**

Zusammenfassung

Eine Vorrichtung zum Analysieren eines Audiosignals hinsichtlich von Rhythmusinformationen des Audiosignals unter Verwendung einer Autokorrelationsfunktion umfaßt eine Filterbank zum Zerlegen des Audiosignals in zumindest zwei Teilbandsignale. Die Teilbandsignale werden hinsichtlich von Periodizitäten mittels einer Autokorrelationsfunktion untersucht, um Rhythmus-Rohinformationen für die zumindest zwei Teilbandsignale zu erhalten. Um die Mehrdeutigkeiten der Autokorrelationsfunktion für periodische Signale zu verringern oder zu eliminieren, werden die Rhythmus-Rohinformationen nachbearbeitet, um nachbearbeitete Rhythmus-Rohinformationen für das Teilbandsignal zu erhalten. Die Rhythmusinformationen des Audiosignals werden auf der Basis der nachbearbeiteten Rhythmus-Rohinformationen ermittelt. Durch die teilbandweise AKF-Nachverarbeitung werden AKF-Mehrdeutigkeiten bereits dort, wo sie entstehen, eliminiert, bzw. werden Rhythmusanteile bei doppelten Tempi, die eine Autokorrelationsfunktionsverarbeitung normalerweise nicht liefert, hinzugefügt, so daß sich als Resultat eine robustere Bestimmung der Rhythmus-Informationen des Audiosignals ergibt.

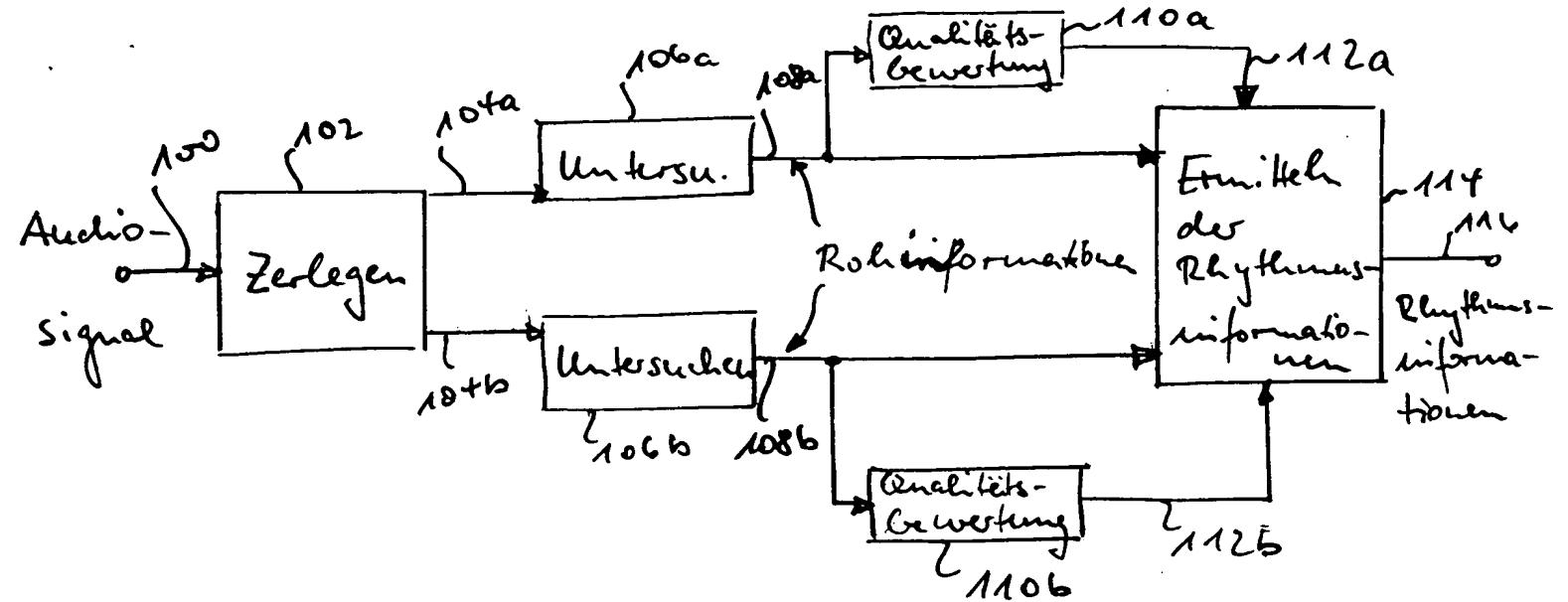


Fig. 1

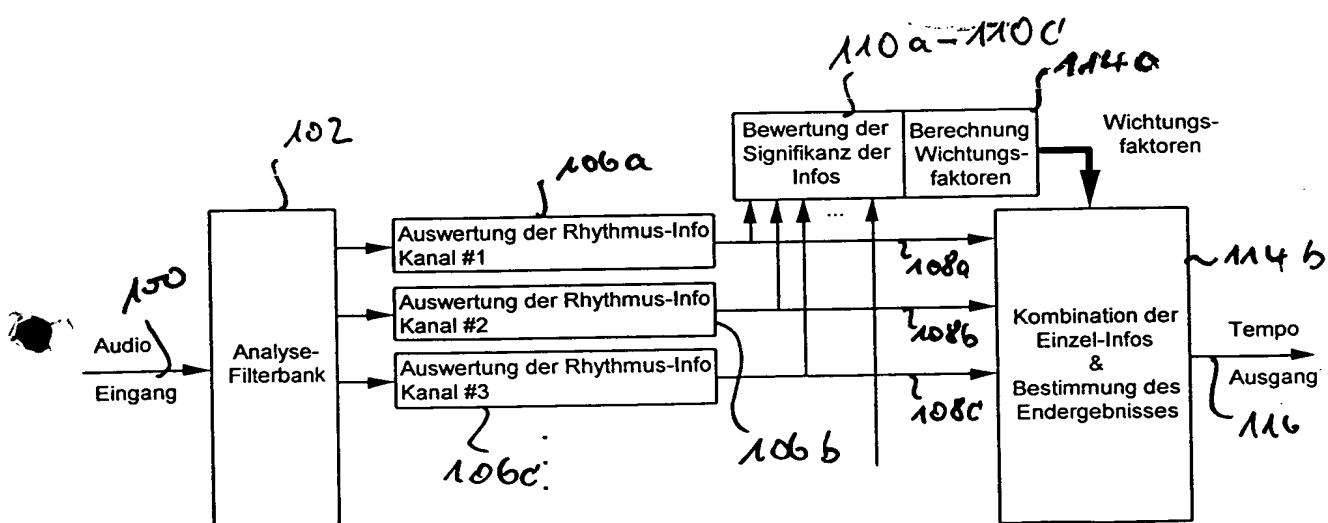


Fig. 2

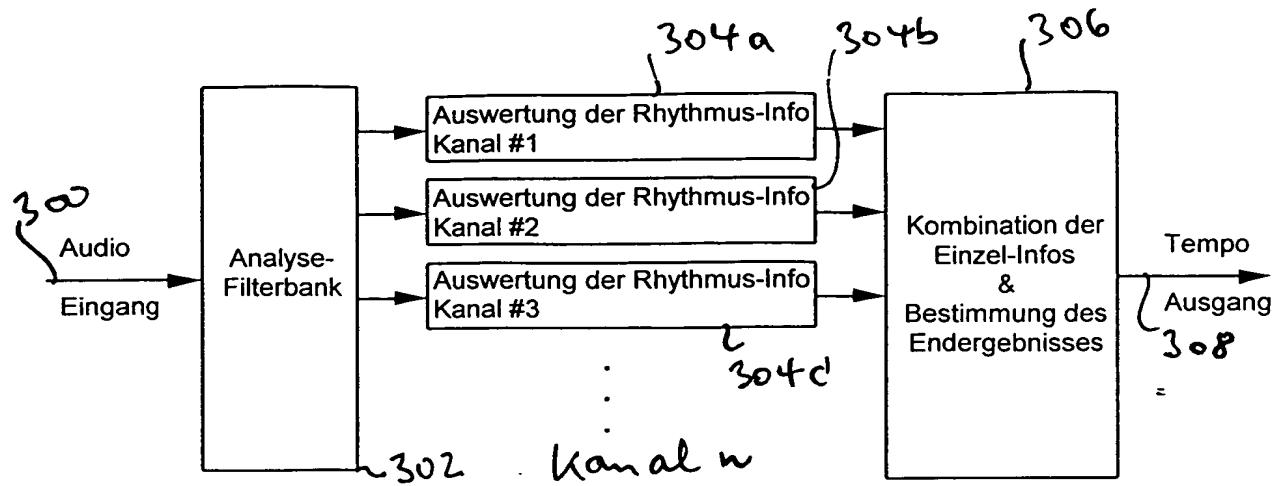


Fig. 3 (Stand der Technik)

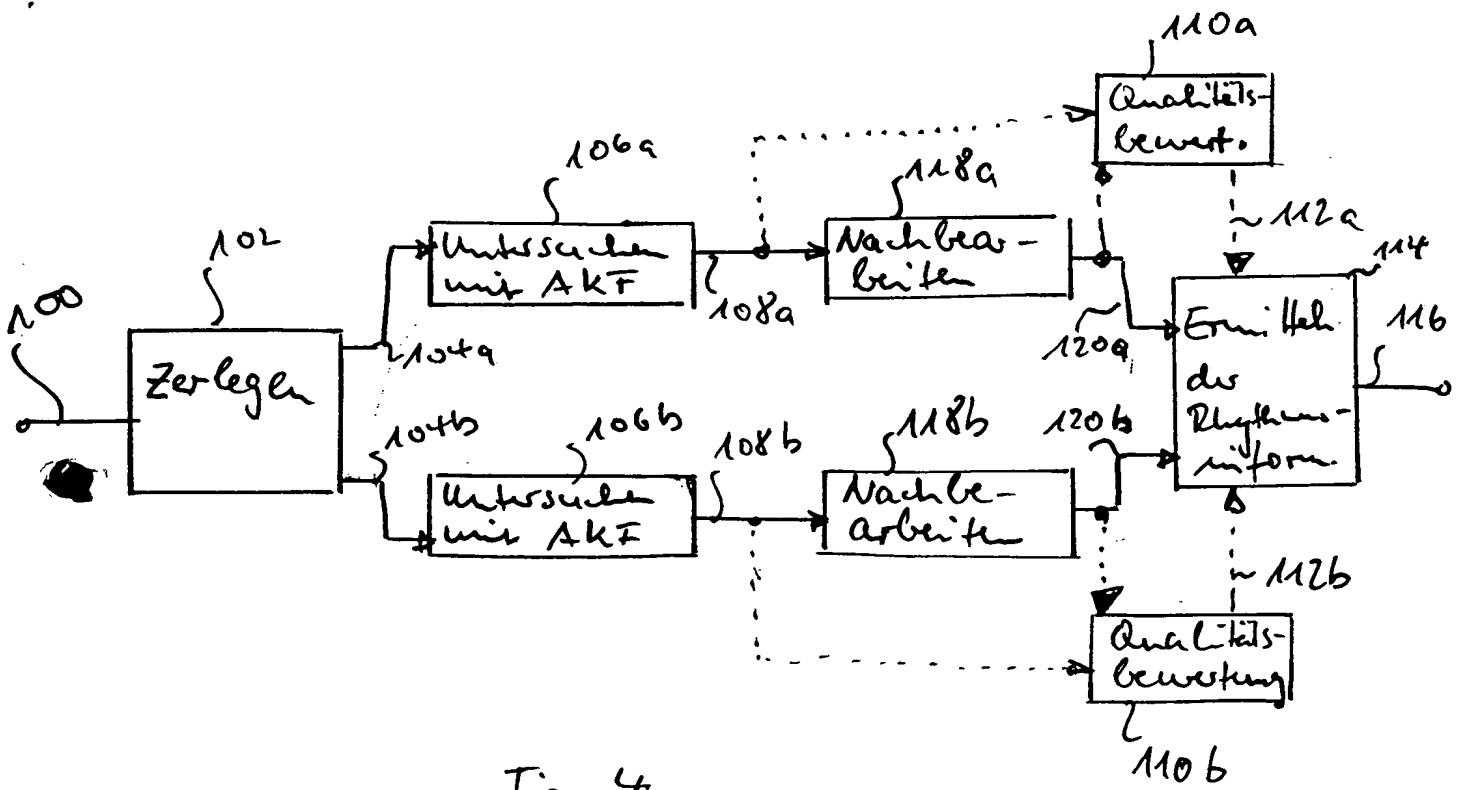


Fig. 4

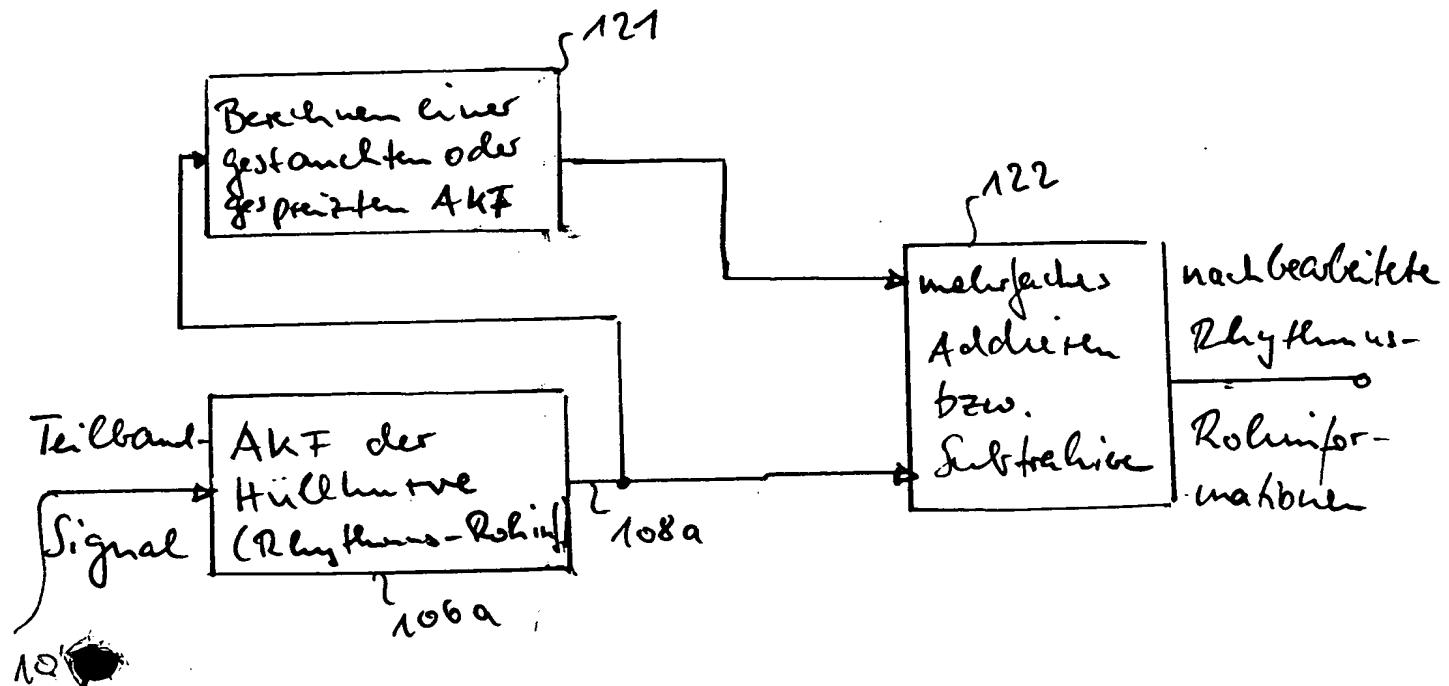


Fig. 5



Creation date: 02-11-2004

Indexing Officer: MVUONG - MY-HANH VUONG

Team: OIPEBackFileIndexing

Dossier: 10659719

Legal Date: 01-21-2004

No.	Doccode	Number of pages
1	CFILE	2
2	SPEC	9
3	CLM	4
4	ABST	1

Total number of pages: 16

Remarks:

Order of re-scan issued on .....